

# PSYCHOLOGY INFORMATION FOR STUDENTS

NO.2 - SEPTEMBER 2004

1. What Really Happened - Bowlby's Forty-Four Juvenile Thieves
2. Experimenter Effects: The Power of Expectations
3. Problems with Research into Gender Differences

ISSN: 1743-3851

Orsett Psychological Services  
PO Box 179  
Grays  
Essex  
RM16 3EW

[orsettpsychologicalservices@phonecoop.coop](mailto:orsettpsychologicalservices@phonecoop.coop)

## Contents

### Page Number

1. What Really Happened - Bowlby's Forty-Four Juvenile Thieves	3
2. Experimenter Effects: The Power of Expectations	12
3. Problems with Research into Gender Differences	18
4. Past Issues	26

# What Really Happened - Bowlby's Forty-Four Juveniles Thieves

## INTRODUCTION

John Bowlby is best known for his work on the effects of maternal deprivation in the early years on later development. His life's work is brought together in the three volumes of "Attachment and Loss" (Bowlby 1969; 1973; 1980), but his original work is known as the "Forty-Four Juvenile Thieves" study (Bowlby 1944a + b) (1)(2).

## BACKGROUND

Juvenile theft accounted for nearly half of all property crime in the early part of the 20th century (Dixon 2003). Pre-World War II thinking on criminality was dominated by the 19th century work of Lombroso (1876) (3), who saw such behaviour as innate. However, psychoanalytic views (4) were growing in popularity - not only to explain the childhood origins of adult behaviour, but also to allow for the possibility of change (Dixon 2003).

John Bowlby completed his medical studies at the University of Cambridge, and also received training in Kleinian object-relations psychoanalysis (5) in the 1920s.

In an early placement, Bowlby came across an isolated teenager who had been expelled from school for theft. He had had no stable mother figure (Bretherton 1992). This experience, and others similar, at a school for "maladjusted children" influenced Bowlby.

## FORTY-FOUR JUVENILE THIEVES - METHOD

This study is based around 44 "juvenile thieves" and 44 non-delinquent controls interviewed at the London Child Guidance Clinic between 1936 and 1939 (6). Initially all children were interviewed with their mothers, where alive, and also given an intelligence test (7). The IQ ranged from 68-159 on the Binet scale with one third of the total sample with an IQ over 130 (8).

The "juvenile thieves" were classified into four groups of stealing based on frequency and duration of behaviour; eg: Grade IV = "persistent thievery over one or three years". Over half of the children (23) were rated as Grade IV compared to four children rated as Grade I ("single theft").

The ages ranged from five years to 16 years 11

months, and the majority (31) were boys <sup>(9)(10)</sup>. The non-delinquent control group were of similar age and intelligence <sup>(11)</sup>, and had been referred to the clinic for reasons other than stealing.

## FINDINGS

Bowlby distinguished six character types among the sample:

- A. "Normal"
- B.(1) Depressed  
(2) Priggish - symptoms of anxiety or hysteria
- C. Circular - alternating depression and over-activity
- D. Hyperthymic - constantly over-active
- E. Affectionless - "characterised by lack of normal affection, shame or sense of responsibility"
- F. Schizoid - schizoid or schizophrenic symptoms.

The "affectionless" character (type E) was the most interesting - shown by 14 of the "juvenile thieves" and none of the control group (table 1). Furthermore, 13 of the fourteen were rated as Grade IV (most severe) stealing <sup>(12)</sup>.

CHARACTER TYPE	THIEVES	CONTROLS
A	2	3
B1	9	13
B2	0	8
C	2	1
D	13	0
E	14	0
F	4	9

(After Dixon 2003)

Table 1 - Number of "juvenile thieves" and controls by character type.

Bowlby wanted to find the cause of the behaviour of the "juvenile thieves" <sup>(13)</sup>. There was little evidence for genetic differences, as much as could be established at that time. Forty-four percent of each group had psychopathology among their parents or grandparents.

Bowlby was more interested in the early experiences, which became the basis of the "maternal deprivation hypothesis" <sup>(14)</sup>. In particular, prolonged separation from the mother, or "broken home", and negative emotional attitudes of the parent(s) towards their child.

i) "Broken home", or prolonged separation in the first five years of life due to mother's death, illness, or hospitalisation.

This type of event was evident in 40% (17) of the "juvenile thieves", and 2% of the controls, but in 85% (12/14) of the "affectionless" character type <sup>(15)</sup>.

Bowlby had "little doubt that prolonged mother-child separations are associated to a high degree both with chronic delinquency in general and with certain types of chronic delinquency in particular" (1944b p110).

Bowlby illustrated this findings with the case of "Derek B", who got diphtheria at 18 months old. He was hospitalised for nine months, during which time his parents did not visit. On returning home, he acted like a "little stranger", preferred to be alone, and later showed problem behaviour at school. This is an example of the disruption of an existing attachment bond, which Bowlby saw as "pathogenic" <sup>(16)</sup>.

ii) Negative emotional attitudes of parents towards their child.

Bowlby noted that a large number of the sample did not have prolonged separation, but their problems (both stealing and otherwise in the control group) were caused by the parent(s)' behaviour. For example, the child being unwanted or disliked by the parent(s).

Mothers with "unbalanced" attitudes towards their children occurred in 76% (32/42) of controls, and 71% (17/24) of "juvenile thieves" not separated from their mothers. For example, "Fred B": "mother very anxious and critical, shouts and terrifies child" <sup>(17)</sup>.

Bowlby also found a number of other relationships between the early experiences and the character type:

- Ambivalent mother and/or recent trauma (eg: death of mother), and Depressed character;
- Ambivalent mother and Hyperthymic character.

Overall, the most striking pattern was the difference between Grade IV stealers and the other "juvenile thieves" and the controls. The former group were much more likely to have suffered prolonged separation from their mothers (table 2) <sup>(18)</sup>.

DEGREE OF STEALING	NUMBER	% SHOWING PROLONGED SEPARATION
Grade IV	23	61
Grade I,II,III	21	14
Controls	44	5

(After Dixon 2003)

Table 2 - Amount of prolonged separation from mothers based on degree of stealing.

## CONCLUSION

Dixon (2003) is positive about the importance of the "juvenile thieves" study as the origin of Bowlby's ideas:

Bowlby's work as a whole was a major contribution to academic thinking about the development of attachment and affectional bonds, and the consequences of their disruption. Bowlby demonstrated that attachment of the infant to the mother is of over-whelming importance in determining the individual's later security and success in forming relations with others, and that separation from or loss of the mother can have a devastating effect  
(p288) (19).

## EVALUATIVE NOTES

1. "The significance of 'forty-four juvenile thieves' as a precursor to attachment theory cannot be overstated" (Dixon 2003 p286).
2. Though Bowlby's work has been controversial, it has provoked further research, over half a century, into the role of attachment in children's development.
3. Lombroso argued that criminals were born different to the rest of the population. What he called "homo delinquens" with underlying genetic differences. Criminals were primitive genetic forms with large jaws and high cheekbones, for example (Brewer 2000).
4. Psychoanalysis is based upon the work of Sigmund Freud (eg: 1940). There are a number of assumptions about behaviour and psychological development. The most relevant here is the importance of early experiences for adult development:

Internal versions of the external world from very early in life. They are not one-to-one mappings of the external world, nor of experience in it, but are emotionally charged constructions  
(Thomas 1996 p290).

5. There are a number of differences between Freudian views, and those of Melanie Klein (eg: 1959) and object-relations theory (eg: Winnicott 1964). A key difference is whether the infant is seeking pleasure and thus motivated to interact with the mother (Freudian view), or that the infant is a "people-seeker", and that is the motivation for interactions.

6. This study is not an experiment because there was not random sampling of participants, which is one of the requirements of the "true" experiment. A "true" experiment, as opposed to a "quasi-experimental" design has randomisation of participants, standardised procedure, and control over the variables (Brewer 2002).

Bowlby also knew which group was which, leading to possible bias (Moxon et al 2003).

"Forty-four juvenile thieves" is a clinical case study, which has both advantages and disadvantages (table 3) (Brewer 2002).

#### ADVANTAGES

- builds up detailed picture of participants\*
- not artificial
- source of future hypotheses\*
- useful for treating individual problems
- helps to discover how past influences present\*
- studying unusual may help with usual
- outstanding cases can be studied
- insight can be gained from one example
- tentative support for theory or challenge orthodox view\*

#### DISADVANTAGES

- not possible to generalise findings\*
- often depends on accuracy of memory of participants when interviewed\*
- past events sometimes explained based on current state of mind\*
- cause and effect not possible to establish\*
- criteria for inclusion of information is subjective\*
- situation and time bound
- poor replicability\*
- biased observer\*

(\* = particularly relevant to this study)

Table 3 - Advantages and disadvantages of the case study method.

7. The interviews were performed by a psychiatrist and a social worker, and a psychologist tested the children's IQ. Bowlby was not necessarily involved here. Because these individuals were patients at the clinic, there were case conferences, when Bowlby was probably involved. Information was also collected from schoolteachers, and from other relatives.

8. Because the sample was based upon referrals to the clinic by schools, parents, courts, and probation officers, intelligent children with problems may have received more attention (Dixon 2003). Simplistically, problem behaviour is often seen as usual with low intelligence.

9. Socio-economic status was not measured.

10. Stealing behaviour among girls would be viewed differently to that of boys. Females tend to be treated more harshly when they break the rules (Edwards 1984).

A woman who commits a crime is in essence breaking two rules - one the rule of law and the other a rule constructed by society as to how she is expected to behave (Nadel 1993 p135).

11. Exact matching of the groups would be difficult, particularly as the study was based on referrals to the clinic only.

12. Bowlby explained the relationship between severe stealing and the "affectionless" character type using psychoanalytic ideas.

The child feels rage from the separation with the parent, and this inhibits their ability to form relationships. This energy is directed into inflicting suffering on others, for example, in the form of stealing.

The lack of the mother stops the development of the superego, which could control such stealing behaviour. Furthermore, the child's rage is reflected on to themselves and the world, leading them to "expect nothing and to feel they deserve nothing" (Dixon 2003).

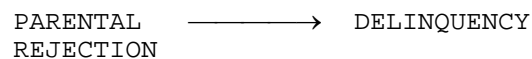
Thus indifference towards others, stops the disappointment of wanting something and not getting it - a "policy of self-protection against the slings and arrows of their own turbulent feelings" (Bowlby 1944b p124).

"Whatever we do", we might imagine them saying, "do not let us care too much for anyone. At all costs let us avoid any risk of allowing our hearts to be broken again" (Bowlby 1944b p124).



13. Recent research (eg: Rutter et al 1998) looks at the correlation between behaviours rather than one behaviour causing another. For example, parental rejection may not cause delinquency, but it is associated with it because of other factors, like the child's temperament (figure 1).

#### CAUSATION



#### CORRELATION

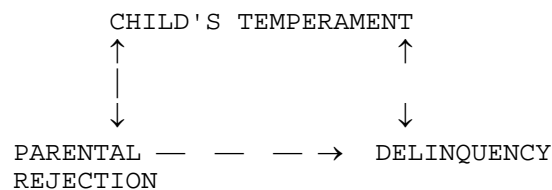


Figure 1 - Difference between causation and correlational relationships.

14. Bowlby proposed the "maternal deprivation hypothesis" in 1951 (developed in 1953 and 1965).

The term "maternal deprivation" is often assumed to be "one specific syndrome of unitary causation". Bowlby meant a number of experiences by this term, and not just "mother-blaming" (Dixon 2003). Ainsworth (1962) attempted to rectify the misunderstandings over this term.

15. A traditional criticism of this study has focused upon the fact that only a minority of "juvenile thieves" had experienced prolonged separation, while the majority had not. "It might seem more logical to conclude that lack of maternal separation leads to delinquency" (Moxon et al 2003 p83).

But the difference is between the number in the "juvenile thieves" group and the control group (40% vs 2%).

16. Six of the 17 "juvenile thieves" with prolonged separation had failed to form a bond at all with the natural mother, as they were separated from her within a few weeks of birth. However, they seemed to form attachments in their foster homes. But serial placements had disrupted this process (Dixon 2003).

17. Bowlby is remembered for his emphasis on the behaviour of the mother, but he did note that five of the "juvenile thieves" (with "normal" mothers) had fathers who never wanted children.

18. Replication of these findings by Bowlby himself at a later date was not found. Children who experienced a period of their early lives in a TB sanatorium were not significantly maladjusted compared to controls in later childhood (Bowlby et al 1956). This study compared 60 children, who spent between six months and two years in hospital before the age of 4 years, with three groups of controls. The children were followed up when 7-14 years old.

Bowlby himself admitted that:

..statements implying that children who are brought up in institutions or who suffer other forms of serious privation and deprivation in early life commonly develop psychopaths or affectionless characters are seen to be mistaken  
..Outcome is immensely varied (Bowlby et al 1956).

19. The main criticism of Bowlby's work was Michael Rutter (1973; 1981).

## REFERENCES

Ainsworth, M.D (1962) The effects of maternal deprivation: A review of findings and controversy in the context of research strategy, Deprivation of Maternal Care, Geneva: World Health Organisation

Bowlby, J (1944a) Forty-four juvenile thieves: Their characters and home-life, International Journal of Psycho-analysis, 25, 19-53

Bowlby, J (1944b) Forty-four juvenile thieves: Their characters and home-life' International Journal of Psycho-analysis, 25, 107-128

Bowlby, J (1951) Maternal Care and Mental Health, Geneva: World Health Organisation

Bowlby, J (1953) Child Care and Growth of Love, London: Tavistock Books

Bowlby, J (1965) Child Care and Growth of Love (new ed), London: Tavistock Books

Bowlby, J (1969) Attachment and Loss vol 1: Attachment, New York: Basic Books

Bowlby, J (1973) Attachment and Loss vol 2: Separation, New York: Basic Books

Bowlby, J (1980) Attachment and Loss vol 3: Loss, Sadness and Depression, New York: Basic Books

Bowlby, J; Ainsworth, M; Boston, M & Rosenbluth, D (1956) The effects of mother-child separation: A follow-up study, British Journal of Medical Psychology, XXIX, 3 & 4, 21-247

Bretherton, I (1992) The origins of attachment theory: John Bowlby and Mary Ainsworth, Developmental Psychology, 28, 759-775

Brewer, K (2000) Psychology and Crime, Oxford: Heinemann

Brewer, K (2002) Research Methods in Clinical Psychology, Psychiatry,

and the Mental Health Professions, Orsett, Essex: Orsett Psychological Services

Dixon, A (2003) "At all costs let us avoid any risk of allowing our hearts to be broken again": A review of John Bowlby's forty-five juvenile thieves, *Clinical Child Psychology and Psychiatry*, 8, 2, 278-289

Edwards, S (1984) *Women on Trial*, Manchester: Manchester University Press

Freud, S (1940) An outline of psychoanalysis. In Strachey, J (ed) (1964) *Standard Edition* vol 23, London: Hogarth Press

Klein, M (1959) Our adult world and its roots in infancy. In Klein, M (1993) *Collected Works* vol III. *Envy, Gratitude and Other Works*, London: Karnac Book

Lombroso, C (1876) *L'Uomo Delinquente*, Milan: Torin

Moxon, D; Brewer, K & Emmerson, P (2003) *Heinemann Psychology AS for AQA A*, Oxford: Heinemann

Nadel, J (1993) *Sara Thornton: The Story of a Woman who Killed*, London: Victor Gollancz

Rutter, M (1973) *Maternal Deprivation Reassessed*, Harmondsworth: Penguin

Rutter, M (1981) *Maternal Deprivation Reassessed* (2nd ed), Harmondsworth: Penguin

Rutter, M; Giller, H & Hagell, A (1998) *Antisocial Behaviour by Young People*, Cambridge: Cambridge University Press

Thomas, K (1996) The defensive self: A psychodynamic perspective. In Stevens, R (ed) *Understanding the Self*, London: Sage

Winnicott, D.W (1964) *The Child, The Family and the Outside World*, Harmondsworth: Penguin

Kevin Brewer

Article written August 2004

## Experimenter Effects: The Power of Expectations

"Experimenter effects" is a general term used to cover inadvertent influences on the participants in an experiment by the experimenter. They are not deliberate, and are done unconsciously by the experimenter (1).

"Experimenter effects" are based around the power of expectations on how participants in each condition of the experiment will behave.

Such effects can have serious consequences for experimental research, particularly if the participant is behaving as expected and not how they would usually behave (2)(3).

The exact mechanism of the influences of expectations on behaviour are not completely clear, but classic work on the process was carried out in the 1960s in two experiments by Rosenthal (4).

### 1. Rosenthal and Fode (1963)

This lab experiment looked at the power of expectations upon animal behaviour.

Twelve psychology students at a US university were asked to train five rats each to run a maze to find food. The students were led to believe that the rats were either specially bred to be good at mazes ("maze-bright" rats) or especially bad ("maze-dull" rats). In fact, both groups of rats were exactly the same.

The rats ran the maze ten times per day for five days. The mean number of correct responses in the maze, and the mean response time to complete the maze were the main measures of the dependent variable.

The "maze-bright" rats were described to the students as expected to learn the maze on the first day, and then get faster each day. While the "maze-dull" rats were expected to show "very little evidence of learning" during the training. The students believed that the purpose of the experiment was to give them experience of training rats for research purposes.

After five days of training, the data showed significant differences between the two groups of rats (table 1). This study showed that students' expectations influenced their behaviour with the rats, and thus the rats behaviour (5).

	MAZE-BRIGHT RATS	MAZE-DULL RATS	
Mean number of correct responses per rat per day	2.32	1.54	(p = 0.01)
Mean time in minutes required to make correct responses	2.35	3.47	(p = 0.02)

(After Rosenthal and Fode 1963)

Table 1 - The differences in maze behaviour between maze-bright and maze-dull rats.

## 2. Rosenthal and Jacobson (1966; 1968)

This field experiment looked at the power of teachers' expectations upon their pupils' educational performance, and is the basis of the "self-fulfilling prophecy" (6).

This is the "idea that expectations about a person or group can become true simply because they have been stated" (Banyard and Grayson 2002).

Based in a US elementary school, this research took place over a school year. At the beginning of the year, 18 classes of children in six grades were given an IQ test (7). From the results of this test, teachers were informed that certain children (20%) were "bloomers" (expected to show rapid intellectual development in the next year). In fact, these individuals were chosen at random.

After eight months, all the children were given the IQ test again to measure their progress.

Over the 8-month period, the growth in IQ score was significantly higher in the experimental group ("bloomers") compared to the rest of the children. For example, 21% of the experimental group showed an increase in IQ score of 30 points or more, compared to only 5% of the control group.

The mean gain was 13.6 points in the experimental group, and 9.0 points for the others. The effect was stronger for younger children (eg: 1st grade: mean increase was 27.4 and 12.0 points respectively) (8).

It is argued that without consciously planning it, teachers were behaving differently towards the "bloomers", and this influenced the children's behaviour leading to their increase in IQ scores.

The teachers were observed to be different towards the "bloomers" in four ways, known as:

i) "climate" - nicer and more friendly to those children (9);

- ii) "input factor" - teach more information to them;
- iii) "response opportunity" - give them greater opportunity to speak in class (and also the length of time allowed to speak);
- iv) "feedback" - more given.

Other ways that teachers can communicate expectations include seating pupils of low-expected ability farther away, and the amount of eye-contact with the pupil (more for high-expected pupils) (10).

The findings of Rosenthal and Jacobson have very important implications, and many studies have attempted to replicate them. Early support showing differences in pupils' performance came from Braun (1976).

But Barber (1976) argued that forty experiments between 1968 and 1976 did not find support. The simplicity of the relationship between teachers' expectations and pupils' educational performance found by Rosenthal and Jacobson is now disputed (Rogers 1998).

Much research has been performed in educational settings to isolate the effects of the teachers' expectations. Rogers (1998), for example, looked at the role of high expectations within the whole school, as well as making distinction between "probabilistic expectations" (most likely to happen) and "prescriptive expectations" (ought to happen).

The pupils' expectations of the teacher can have an effect as well as the teachers' expectations of the pupils (Feldman and Theiss 1982) (11).

#### FOOTNOTES

1. Barber (1976) listed ten pitfalls for experimenters in running experiments (table 2), and called the effect of interest here, "experimenter unintentional expectancy effect".

2. One way to overcome "experimenter effects" is through the double-blind design. Neither the participants nor the experimenter carrying out the procedure knows which group is which, and what is expected to happen in the experiment (Brewer 2002).

3. Researchers influencing participant's behaviour has also been found in interviews. Greenspoon (1955) reported the "Greenspoon effect", where it is possible to influence the answers given by the use of "mm-hmm" or "uh-huh" as subtle reinforcement.

1. Investigator\* paradigm effect
2. Investigator experimental design effect
3. Investigator loose procedure effect
4. Investigator data analysis effect
5. Investigator fudging effect
6. Experimenter personal attributes effect
7. Experimenter failure to follow the procedure effect
8. Experimenter misrecording effect
9. Experimenter fudging effect
10. Experimenter unintentional expectancy effect

\* Investigator = person in overall control of research; experimenter = person carrying out procedure; often same person

(After Coolican 1990)

Table 2 - Ten pitfalls of running experiments by Barber (1976).

4. See also Rosenthal (1966).

5. Barber (1976) argued that these findings are more likely to be due to "experimenter failure to follow the procedure effect".

6. In a classic example, Guthrie (1938) reported a case where a group of boys decided to treat a physically unattractive girl as if she was physically attractive. They each asked to take her out on a date. Soon she was perceived as physically attractive because her self image had changed leading to her behaving differently.

In another example, the Ashanti people in West Africa name their children based on the day of the week born. Associated with each day of the week are characteristics; eg: Wednesday-born boys are believed to be aggressive, while Monday-born boys are quiet.

Jahoda (1958) noted that there are disproportionately more Wednesday-born boys in police records. In other words, they are expected to be aggressive, and this expectation comes true.

The ideas of the self-fulfilling prophecy are more formally adopted in labelling theory (Scheff 1966), and symbolic interactionism (Blumer 1969), and by Becker (1963) and Lemert (1972) to explain deviant behaviour.

The self-fulfilling prophecy is also involved in forming impressions of others. It works as follows: person A believes person B is a certain type of person (for example, through stereotyping); person A then acts towards B based on that belief; person B responds to that behaviour; this confirms person A's belief about B.

7. Tests of General Ability (TOGA) (Flanagan 1960 quoted in Banyard and Grayson 2002).
8. It has been suggested that the IQ test used was not age-standardised, and that the gains in IQ were a facet of the test rather than real improvements due to expectations (Banyard and Grayson 2002). The increases in IQ scores over a short period were in some cases exceptionally large (eg 30 pts in 8 months).
9. Early research by Brophy and Good (1970) confirmed more teacher praise to high-expectation pupils, and these children are also given more attention (Rothbart et al 1971).
10. Experimenters can influence participant's behaviour through non-verbal behaviour (eg: facial expressions) and verbal cues (eg: tone of voice). Barber (1976) called these, "experimenter personal attributes effect".
11. The power of expectations is also seen in the "placebo effect" related to health (Harrington 2000; Walach 2002).

## REFERENCES

- Banyard, P & Grayson, A (2002) *Introducing Psychological Research* (2nd ed), Basingstoke: Palgrave
- Barber, T.X (1976) *Pitfalls in Human Research*, Oxford: Pergamon
- Becker, H (1963) *Outsiders*, New York: Free Press
- Blumer, H (1969) *Symbolic Interactionism*, Englewood Cliffs: Prentice-Hall
- Braun, C (1976) Teacher expectation: sociopsychological dynamics, *Review of Educational Research*, 81, 74-84
- Brewer, K (2002) *Research Methods in Clinical Psychology, Psychiatry, and the Mental Health Professions*, Orsett, Essex: Orsett Psychological Services
- Brophy, J.E & Good, T.L (1970) Teacher's communication of differential expectations for children's classroom performance: some behavioural data, *Journal of Educational Psychology*, 61, 365-374
- Coolican, H (1990) *Research Methods and Statistics in Psychology*, London: Hodder & Stoughton
- Feldman, R.S & Theiss, A.J (1982) The teacher and student as pygmalsions: Joint effects of teacher and student expectations, *Journal of Educational Psychology*, 74, 2, 217-223
- Greenspoon, J (1955) The reinforcing effect of two spoken sounds on the frequency of two responses, *American Journal of Psychology*, 68, 409-416
- Guthrie, E.R (1938) *Psychology of Human Conflict*, New York: Harper
- Harrington, A (2000) The whiteness of lies: Swallowing the placebo effect, *Cerebrum*, 2, 1, 71-86



Jahoda, M (1958) Current Concepts of Positive Mental Health, New York: Basic Books

Lemert, E.M (1972) Human Deviance, Social Problems and Social Control (2nd ed), Englewood Cliffs: Prentice-Hall

Rogers, C (1998) Teacher expectations: Implications for school improvement. In Shorrock-Taylor, D (ed) New Directions in Educational Psychology, London: Routledge

Rosenthal, R (1966) Experimenter Effects in Behavioural Research, New York: Appleton-Century-Crofts

Rosenthal, R & Fode, K.L (1963) The effect of experimenter bias on the performance of the albino rat, Behavioural Science, 8, 183-189

Rosenthal, R & Jacobson, L (1966) Teachers' expectancies: Determinants of pupils' IQ gains, Psychological Reports, 19, 115-118

Rosenthal, R & Jacobson, L (1968) Pygmalion in the Classroom, New York: Holt, Rinehart & Winston

Rothbart, M; Dalfen, S & Barrett, S (1971) Effects of teacher's expectancy on student-teacher interaction, Journal of Educational Psychology, 62, 49-54

Scheff, T.J (1966) Being Mentally Ill: A Sociological Theory, Chicago: Aldine

Walach, H (2003) Placebo and placebo effects - a concise review, Focus on Alternative and Complementary Therapies, 8, 2, 178-187

Kevin Brewer

Article written August 2004

## Problems with Research into Gender Differences

The idea that there are differences between men and women in certain behaviours is accepted as commonsense (eg: "Men are from Mars, Women are from Venus"; Gray 1993). These differences are confirmed by regular reports of psychological research finding such differences.

The type of research used to study gender differences are based around certain methodologies including:

### i) Self-rated inventories or questionnaires

Here participants rate themselves on adjectives, say, to describe their behaviour. For example, the Williams and Best (1982; 1990) fourteen country study found that men rated themselves more as "ambitious" and "dominant" than women, and women as "sensitive" and "kind" compared to men.

However, this method does assume that the self reports are accurate (Brewer 2002).

### ii) Personality inventories or tests

Research using standardised and validated personal inventories or tests show differences between men and women like men scoring higher on characteristics of "assertiveness", and women on "tendermindedness" (Feingold 1994).

This method has both advantages and disadvantages (table 1).

ADVANTAGES	DISADVANTAGES
- individual's data can be compared to other individuals	- depends on accuracy of construction
- individual's data can be compared to norms	- assumes certain norms
- standardised comparison between studies	- depends on honesty of answers
	- need co-operation of test-taker
	- can be faked
	- measuring "constructs" not "existing things"

Table 1 - Advantages and disadvantages of personality inventories.

### iii) Summaries of research/literature review

Maccoby and Jacklin (1974) produced the most comprehensive survey of research on gender differences from over 2000 books and articles. They grouped the findings into three categories:

a) myths of gender differences; eg: girls are more "social" than boys;

b) evidence of differences; eg: males are more aggressive than females;

c) ambiguous or inconclusive studies; eg: differences in tactile sensitivity.

This research simply counted the number of findings, and all studies were given equal weight, even if there were methodological weaknesses. Thus there is a need for a more sophisticated way of summarising the research, which came with meta-analysis.

### iv) Meta-analysis

This technique summarises other methods of research by statistically reanalysing them, and producing an overall score of difference or "effect size" (Wood 1995). Findings here included that men exceed women on physical aggression ( $d = .40$ ; Eagly and Steffen 1986) and restlessness ( $d = .72$ ; Eagly 1987), and vice versa for measures of conversational sensitivity (eg: decoding visual cues  $d = -.32$ ; auditory cues  $d = -.18$ ; Hall 1984).

Meta-analysis is based on the comparison of group means using the "d" statistic <sup>(1)</sup>. The "d" statistic is expressed in standard deviation units, which is assuming a normal distribution of behaviour. It is possible that behaviour for men or women is not distributed in such a way. Furthermore, Hayes (1994) argued that cultural and social pressures can skew any distribution in one direction.

Table 2 compares the literature review/summary of studies with meta-analysis.

LITERATURE REVIEW	META-ANALYSIS
- qualitative	- quantitative
- selected coverage	- comprehensive coverage
- individual studies	- overall magnitude of change
either statistically	possible to calculate
significant or not	

Table 2 - Comparison of meta-analysis with traditional literature review.

## GENERAL PROBLEMS

### 1. Published or unpublished studies

Studies finding differences are more likely to be published (known as the "file-drawer problem; Smith and Egger 1998). This leads to a possible over-representation of gender differences. To every one study published showing gender differences, how many unpublished studies did not find differences?

The focus upon difference emphasises the message that there are differences: sameness is of little interest (James 1997).

### 2. Use of average scores

The analysis of differences is based around average scores. James (1997) called this the "tyranny of averages". There will be large individual variations in scores which produce the average. Focusing upon the average may give a false picture of the gender differences.

For example, Hyde (1986) found a medium difference in mathematics ability between men and women, but, at the extreme of high ability, there were five times more males.

### 3. Method used and results

Different types of studies find differing results. Eisenberg and Lennon (1983) looked at the studies on gender differences in empathy. They found that studies using self-reported scales (overt measure of behaviour) had large differences, while lab-based studies moderate differences, and physiological measures (hidden measure) produced no gender differences.

It was suggested that in some situations "demand characteristics" <sup>(2)</sup> (Orne 1992) were involved, and women were likely to live up to the stereotype of being more empathetic than men.

### 4. Underlying assumptions

The underlying assumption of research into gender differences is that:

..it has tended to make assumptions about sex-stereotypical behaviour, without really taking into account issues of similarity or androgyny. The problem is that research which is designed to uncover differences between groups of people

almost always prevents researchers from identifying similarities..(Hayes 1994 p757).

It is implicit in the research that there are two genders, which are different. Bem (1975) has argued for a middle ground known as "androgyny", which is a combination of both masculine and feminine.

## 5. Interpretation of results

If the differences between men and women are accepted as real, are they caused by underlying biological differences (eg: sexual selection to explain differences in mathematics; Geary 1996), or the product of social learning (eg: Eagly 1987: social role theory)?

This is not a simple question because as Raven and Rubin (1983) noted: "behaviour differences between men and women must be considered in the larger context of a society that treats men and women differently".

The study of gender differences exists in a patriarchal <sup>(3)</sup> context (Hollway 1989) with power relations in how men and women relate to each other (Nicolson 1997). It is possible to talk of the "battle of the sexes" discourse <sup>(4)</sup> which underpins much discussion of gender differences. In other words, there are vested interests which want to encourage the emphasis on gender differences <sup>(5)</sup>.

## 6. Differences over time

Some gender differences may be diminishing over time. For example, Maccoby and Jacklin in 1974 found evidence that girls had greater verbal ability, and boys excelled in mathematics ability. By 1988, Hyde and Linn argued the research showed that verbal ability had increased for boys. While, based on GCSE results, girls outperformed boys in mathematics in 2000 (Eysenck and Flanagan 2001).

## 7. The categories of behaviour used.

Some of the categories of behaviour used in the studies are more concrete (eg: aggression) and reliable than others. Women score higher on categories of behaviour like "social sensitivity", which includes role-taking, nurturance, and empathy (Lippa 1994). This is a more subjective category of behaviour. It is much easier (in terms of reliability and validity) to measure

"personal space distance" (which men score higher; Hall 1984) in the research, for example, than "socio-emotional group behaviours" (which women score higher; Carli 1982 quoted in Campbell 1996).

## CONCLUSIONS

In reality, women and men are more alike than different: "there is greater variability between individual women and between individual men than there is between women as a group and men as a group" (Griffin 1991).

The tendency in much research into gender differences was to emphasis girls' poorer performance (Hollway et al 2002), which led feminist psychologists to be critical of poor scientific practice. Hyde (1994) suggested a set of guidelines for non-sexist research on gender differences. It included the recommendations that:

a) Research that found no significant gender differences should also be published;

b) Published work should specify the size of any gender differences;

c) Interpretations of differences should not be based upon the assumption of the male standard as the norm (6);

d) Biological differences should not be inferred from behavioural tests, only from biological measures.

## FOOTNOTES

1. Campbell (1996) defined "d" as the mean of group A divided by the standard deviation of A minus the mean of group B divided by the standard deviation of B. While Glass et al (1981) used the mean of experimental group minus mean of control group, then divided by standard deviation of control group.

2. "Demand characteristics" are where participants in research behave as they feel they ought or how they think the experimenter wants rather than how they usually behave.

3. "Patriarchy.. means the manifestation and institutionalisation of male dominance over women and children in the family and the extension of male dominance over women in society in general. It implies that men hold power in all the important institutions of

society and that women are deprived of access to such power. It does not imply that women are either totally powerless or totally deprived of rights, influence, and resources" (Lerner 1986 quoted in Wooley 1994).

4. "Discourses" are not easy to define. Parker (1992) saw them as a "system of statements which construct an object", and Potter and Wetherell (1987) called them "historically developing linguistic practices". Generally, discourses can be seen as "spoken interactions, written texts, and ideas" (Brewer 2001).

5. The vested interests include those groups or individuals who benefit from such a situation. Within a patriarchal context, the emphasis on difference allows men to distinguish themselves from women, and maintain their dominant position in society.

Connell (1995) talked about a "patriarchal dividend" that men generally gain in terms of honour and prestige, and materially with higher average wages than women. For example, the average full-time income of women is about 80% that of men for comparable jobs (Drever et al 2000).

6. Broverman et al (1970) found that mental health professionals asked to describe a mentally healthy adult, man, and woman tended to equate "what it was to be a man with what it was to be an adult. Men were taken as the human norm and women were contrasted to this" (Hollway et al 2002 p120).

## REFERENCES

Bem, S.R (1975) Sex role adaptability: One consequence of psychological androgeny, *Journal of Personality and Social Psychology*, 31, 634-643

Brewer, K (2001) An introduction to the social construction of eating disorders among women, *Orsett Psychological Review*, 1, March, 2-10

Brewer, K (2002) *Research Methods in Clinical Psychology, Psychiatry, and the Mental Health Professions*, Orsett, Essex: Orsett Psychological Services

Broverman, J.K et al (1970) Sex-role stereotypes and clinical judgments of mental health, *Journal of Consulting and Clinical Psychology*, 34, 1-7

Campbell, A (1996) Sex differences and social behaviour in children and adults. In Hatcher, D (eds) 1996 ATP Conference, Leicester: Association for Teaching of Psychology

Connell, R.W (1995) *Masculinities*, Oxford: Polity Press

Drever, F; Fisher, K; Brown, J & Clark, J (2000) *Social Inequalities*, London: Office for National Statistics

Eagly, A (1987) *Sex Differences in Social Behaviour: A Social Role Interpretation*, Hillsdale, NJ: Erlbaum

Eagly, A & Steffen, V (1986) Gender and aggressive behaviour: A meta-analytic review of the social psychological literature, *Psychological*

Bulletin, 100, 309-330

Eisenberg, N & Lennon, R (1983) Sex differences in empathy and related capacities, *Psychological Bulletin*, 94, 100-131

Eysenck, M & Flanagan, C (2001) *Psychology for A2 Level*, Hove, East Sussex: Psychology Press

Feingold, A (1994) Gender differences in personality: A meta-analysis, *Psychological Bulletin*, 116, 429-456

Geary, D.C (1996) Sexual selection and sex differences in mathematics ability, *Behaviour and Brain Sciences*, 19.2, 229-246

Glass, G; McGraw, B & Smith, M.L (1981) *Meta-analysis in Social Research*, Beverly Hills, CA: Sage

Gray, J (1993) *Men are from Mars, Women are from Venus*, London: Thorson

Griffin, C (1991) Sex differences and gender relations: "Yes, but who does the washing up?" In Cochrane, R & Carroll, D (eds) *Psychology and Social Issues*, Sussex: Falmer Press

Hall, J (1984) *Non-Verbal Sex Differences: Communication Accuracy and Expressive Style*, Baltimore: John Hopkins University Press

Hayes, N (1994) *Foundations of Psychology*, London: Routledge

Hollway, W (1989) *Subjectivity and Method in Psychology*, London: Sage

Hollway, W; Cooper, T; Johnston, A & Stevens, R (2002) The psychology of sex and gender. In Cooper, T & Roth, I (eds) *Challenging Psychological Issues*, Milton Keynes: Open University

Hyde, J.S (1986) Gender differences in aggression. In Hyde, J.S & Linn, M.C (eds) *The Psychology of Gender: Advances Through Meta-analysis*, Baltimore: John Hopkins University Press

Hyde, J.S (1994) Should psychologists study sex differences? Yes, with some guidelines, *Feminism and Psychology*, 4, 4

Hyde, J.S & Linn, M.C (1988) Gender differences in verbal ability: A meta-analysis, *Psychological Bulletin*, 104, 53-69

James, L (1997) (ed) Significance of gender: Theory and research about difference, *Journal of Social Issues*, 53, 2, special issue

Lippa, R (1994) *Introduction to Social Psychology* (2nd ed), Pacific Grove, CA: Brooks/Cole

Maccoby, E.E & Jacklin, C.N (1974) *The Psychology of Sex Differences*, Stanford, CA: Stanford University Press

Nicolson, P (1997) Gender and psychology: Adding gender to the curriculum, *Psychology Teaching*, Oct, 13-18

Orne, M.T (1962) On the social psychology of the psychology experiment: With particular reference to demand characteristics and their implications, *American Psychologist*, 17, 776-783

Parker, I (1992) Discourse discourse: Social psychology and modernity. In Doherty, J; Graham, E & Malik, M (eds) *Postmodernism and Social Sciences*, Basingstoke: Macmillan

Potter, J & Wetherell, M (1987) *Discourse and Social Psychology: Beyond Attitudes and Behaviour*, London: Sage

Raven, B.H & Rubin, J.Z (1983) *Social Psychology* (2nd ed), New York: John Wiley

Smith, G.D & Egger, M (1998) Meta-analysis: Unresolved issues and future developments, *British Medical Journal*, 17/1, 221-225



Williams, J & Best, D (1982) Measuring Sex Stereotypes: A Thirty Nation Study, Beverly Hills, CA: Sage

Williams, J & Best, D (1990) Sex and the Psyche: Gender Roles and Self Concepts Viewed Cross-Culturally, Beverly Hills, CA; Sage

Wood, P (1995) Meta-analysis. In Breakwell, G.M et al (eds) Research Methods in Psychology, London: Sage

Wooley, O.W (1994) And man created "woman". In Fallon, P; Katzman, M.A & Wooley, S.C (eds) Feminist Perspectives on Eating Disorders, New York: Guilford Press

Kevin Brewer

Article written August 2004

ORSETT PSYCHOLOGICAL SERVICES  
PO Box 179, Grays, Essex RM16 3EW

PSYCHOLOGY INFORMATION FOR STUDENTS

ISSN: 1743-3851

Published twice per year: March/September

Cost: £3.00 each (cheques payable to "Kevin Brewer")

No.1: March 2004

- What really happened - the "Bobo" doll experiments
- Contribution of cognitive neuropsychology to understanding cognitive processes
- Ethical issues in non-experimental research
- Hegemony (Gramsci)

No.2: September 2004

- What really happened - Bowlby's forty-four juvenile thieves
- Experimenter effects - the power of expectations
- Problems with research into gender differences